

Broadcasting colourings on trees.

A combinatorial view.*

Charilaos Efthymiou

University of Warwick, Mathematics and Computer Science, Coventry CV4 7AL, UK
c.efthymiou@warwick.ac.uk

Abstract

The broadcasting models on a d -ary tree T arise in many contexts such as discrete mathematics, biology, information theory, statistical physics and computer science. In this paper we consider the k -colouring model where we assign a colour to the root of T and the remaining vertices are coloured as follows: For each child of a vertex v coloured σ_v , assign it a colour chosen uniformly at random among all the colours but σ_v . A basic question here is whether the information of the assignment of the root affects the distribution of the colourings of the leaves. This is the so-called *reconstruction/non-reconstruction problem*. It is well known that $d/\ln d$ is a threshold in the sense that

- if $k > (1 + \epsilon)d/\ln d$, then the information of the colouring of the root vanishes as the height of the tree grows and it does not affect the distribution of the colourings of the leaves.
- if $k < (1 - \epsilon)d/\ln d$, then the information of the colouring of the root affects the distribution of the colouring of the leaves regardless of the height of the tree.

Despite this threshold, there is no apparent *combinatorial* reason why such a result should be true. Searching for an explanation, we note that the above implies that for $k > (1 + \epsilon)d/\ln d$ and two broadcasting processes that assign the root different colours the following holds: There is a coupling such that the probability of having vertices which take different colour assignments in the two processes reduces as we move away from the root.

The description of such a coupling, especially a combinatorial one, is a stronger statement than the reconstruction/non-reconstruction threshold. As it would provide evidence on *why* we have this very impressive phenomenon. Devising such a coupling calls for understanding a complex combinatorial problem and it is a non-trivial task to achieve for any $k < d$.

In this work we provide a coupling which has the aforementioned property for any $k > 2d/\ln d$. Interestingly enough, the decisions that it makes are based on *local rules*, i.e. it considers only two successive levels of the tree. It should not be taken for granted that such a coupling is local for any k down to $d/\ln d$.

Finally, we relate the existence of such a coupling to sampling k -colourings of sparse random graphs, with expected degree d , for $k < d$.

1 Introduction

The broadcasting models and the closely related reconstruction problems on trees were originally studied in statistical physics. Since then they have found applications in other areas including biology (in phylogenetic reconstruction [5, 16]), communication theory (in the study of noisy computation [8]). Very impressively, these models arise in the study of random constraints satisfaction problems such as random k -SAT, random graph colouring etc. That is, these models on trees seem to capture some of the most fundamental properties of the corresponding models on random (hyper)graphs (see [13]).

*Supported by EPSRC grant EP/G039070/2 and DIMAP.

The most basic problem in the study of broadcasting models is determining the reconstruction/non-reconstruction threshold. I.e. whether the configuration of the root affects the distribution of the configuration of the leaves of the tree. The transition from non-reconstruction to reconstruction can be achieved by adjusting appropriately the parameters of the model. Typically, this transition exhibits threshold behaviour. So far, the main focus of the study was to determine the precise location of the reconstruction/non-reconstruction threshold in various models.

In this work we focus on the colouring model, for which the reconstruction/non-reconstruction threshold is known [17, 19, 20, 4]. We investigate further the phenomenon searching for a natural, i.e. combinatorial, explanation of why the information decays in the non-reconstruction regime. An explanation which has been, somehow, elusive for a number of colours k smaller than the number of children per non-leaf vertex, d . As far as the reconstruction regime is regarded a satisfactory, combinatorial, explanation is already known [17, 19]. Here we provide a combinatorial view of the problem by providing a coupling between two broadcasting processes which work for k well below d , i.e. for $k > 2d/\ln d$. That is, it implies non-reconstruction for $k > 2d/\ln d$. Even though the coupling is not optimal, it reveals a great deal of why we have non-reconstruction on trees. Interestingly enough, the decisions that it makes are based on *local rules*, i.e. it considers only two successive levels of the tree. The reader should not take for granted that such a coupling is fairly local for every k down to $d/\ln d$.

Finally, recent advances in sampling colouring algorithms (see [9]) allow to relate such a coupling to sampling k -colourings of random graphs and regular graph for k smaller than the expected degree d . See further discussion in Section 1.2.

1.1 The model

The broadcasting models on a tree T are models in which information is sent from the root over the edges to the leaves. We assume that the edges represent noisy channels. For some finite set of spins $\Sigma = \{1, 2, \dots, k\}$, a configuration on T is an element of Σ^T , i.e. it is an assignment of a spins of Σ to each vertex of T . The spin of the root r is chosen according to some initial distribution over Σ . The information propagates along the edges of the tree as follows: There is a $k \times k$ stochastic matrix M such that if the vertex v is assigned spin i , then its child u is assigned spin j with probability $M_{i,j}$.

Here we focus on the model which is known as the (proper) k -colouring model (or k -state Potts model at zero temperature). In particular, we assume that the underlying tree T is a complete d -ary tree, it is finite with height h and we let L_h denote the leaves of T . Also, for the matrix M it holds that

$$M_{i,j} = \begin{cases} \frac{1}{k-1} & \text{for } i \neq j \\ 0 & \text{otherwise.} \end{cases}$$

Broadcasting models and more specifically colouring can be considered as Gibbs measures on trees. That is, assuming that the broadcasting process over T starts with root r coloured i , then the k -colouring after the processes has finished is a random k -colouring of T conditional the colouring of r . We let μ_i denote the uniform distribution over the configurations of T conditional that r is assigned i .

Definition 1 We say that a model is reconstructible on a tree T if there exists $i, j \in \Sigma$ for which

$$\lim_{h \rightarrow \infty} \|\mu_i - \mu_j\|_{L_h} > 0,$$

where $\|\cdot\|_{L_h}$ denotes the total variation distance of the projections of μ_i and μ_j on L_h . When the above limit is zero for every i, j , then we say that the model has non-reconstruction.

(Non)-Reconstructibility expresses how information decays along the tree. As a matter of fact, non-reconstruction is equivalent to the mutual information between the colouring of root r and that of L_h going to zero as h grows (see [15]). Also, non-reconstruction is equivalent to Gibbs distribution being

extremal (see [11]). That is, *typical* colourings of the leaves do not provide information for the colouring of the root. The notion of extremality is weaker than the well-known *Gibbs uniqueness* where we require that *every* configuration of the leaves should not provide information for the colouring of the root.

An early work about reconstruction/non-reconstruction problems on tree is that of Kesten and Stigum in [12]. The authors there present a general result for reconstruction, the so called “Kesten-Stigum bound” that states that reconstruction holds when $\lambda^2 d > 1$, where λ is the second largest eigenvalue of M in absolute value. This bound is sharp for a lot of models, e.g. Ising model (see [8]). In [15] it was shown that there are models where the Kesten-Stigum bound is not sharp, e.g. the binary models where M is sufficiently asymmetric or ferromagnetic q -state Potts model with q large. As far as the (proper) k -colouring model is regarded the Kesten-Stigum bound is not tight but we know the reconstruction/non-reconstruction threshold quite precisely. From [17, 19, 20, 4] we derive the following theorem:

Theorem 1 *For fixed $\epsilon > 0$ and sufficiently large d , the following is true for the k -colouring model on a d -ary tree T :*

- *If $k \geq (1 + \epsilon)d / \ln d$, then the model is non-reconstructible.*
- *If $k \leq (1 - \epsilon)d / \ln d$, then the model is reconstructible.*

Remark 1 We should remark that reconstruction bound is from [17, 19] and is based on analysing a simple reconstruction algorithm. As a matter of fact the reconstruction condition there is more precise than that in Theorem 1, i.e. it should hold $d > k[\ln k + \ln \ln k + 1 + o(1)]$.

Remark 2 The non-reconstruction bound is from [20, 4]. The result in [20] provides a very precise condition for non-reconstruction, i.e. $d \leq k[\ln k + \ln \ln k + 1 - \ln 2 - o(1)]$. In [4] the reader can find further interesting results about the problem.

Unfortunately, neither of [20, 4] give explicit evidence on why do we have the phenomenon of information decay on the trees. What is the natural (combinatorial) reason why such a result should be true. Searching for an explanation for the non-reconstruction we observe that Theorem 1 with Coupling Lemma (see [3]) imply the following corollary.

Corollary 1 *For any fixed $\epsilon > 0$, sufficiently large d and $k = (1 + \epsilon)d / \ln d$ the following is true for the k -colourings of the d -ary tree T : Any two broadcasting processes which assign the root different colours can be coupled such that the probability of having vertices that take different colour assignments in the two processes reduces as we move away from the root.*

That is, somehow there is a rule which specifies how someone should correspond the choices of colourings in the first broadcasting process to the choices of the other process such that the probability of having a set of vertices that take different colours reduces as the distance of the set from the root increases. The reader should note that the construction of such a coupling is far from trivial for any $k < d$.

Here we address the problem of constructing a coupling as specified in Corollary 1, based on *local combinatorial* rules. By local we mean that once the first process decides on colouring a *fairly small* set of vertices, then we should be able to know how the other process should colour the same set of vertices. Such a construction is an interplay among the combinatorial structure of the tree, its colourings and the Gibbs distribution of the colourings. In this work we provide the following result:

Main Result: We construct a coupling for two broadcasting processes on a d -ary tree T that is combinatorial, local and has the properties specified in Corollary 1 for $k = (2 + \epsilon)d / \ln d$, for fixed $\epsilon > 0$ and sufficiently large d .

1.2 Further Motivation - Non Reconstruction in Random Graphs & Sampling

It is believed that the non-reconstruction/reconstruction phenomenon determines the dynamic phase transition in the random graph $G_{n,m}$ [13]. Where $G_{n,m}$ denotes the random graph on n vertices and m edges while we let d be the expected degree, i.e. $d = 2m/n$. In this context we take d to be fixed.

The *dynamic phase transition* is related to the geometry of k -colourings of $G_{n,m}$ and it was predicted by statistical physicists in [13], based on ingenious however mathematically non-rigorous arguments. Let us be more specific. For typical instances of $G(n, m)$, it is well known that the chromatic number $\chi \sim \frac{d}{2 \ln d}$ (see [2]). The 1-step Replica Symmetry breaking hypothesis (see [13]) suggests that considering the space of k -colourings of $G_{n,m}$ as k varies from large to small, we will observe the following phenomenon: For $k = (1 + \epsilon)d / \ln d$ (i.e. greater than 2χ) all but a vanishing fraction of k -colourings form a giant connected ball. The notion of connectedness suggests that starting from any colouring we can traverse the whole set of solutions by moving in steps. Each steps involves changing only a very small -constant- number of colour assignments. However, for $k = (1 - \epsilon)d / \ln d$ the set of k -colouring shatters into exponentially many connected balls with each ball containing an exponentially small fraction of all k -colourings. Any two colourings in different balls are separated with linear hamming distance. For a rigorous description of the shattering phenomenon see [1].

We should remark that the dynamic phase transition in $G_{n,m}$ (roughly) coincides with the extremality /non-extremality transition of the colourings in a d -ary tree. Further investigation into this coincidence yield the authors in [10] to developed a sufficient condition for the tree and random graph reconstruction problem to coincide. In [14] this condition was verified for symmetric models like colouring.

It is believed we can have approximate¹ random colouring of $G_{n,m}$ in the whole regime of non-reconstruction. Recently, the author of this paper in [9] suggested a new algorithm for sampling colourings of $G_{n,m}$ with constant expected degree. Interestingly enough the accuracy of the algorithm depends directly on non-reconstruction conditions. The idea there is that we first remove edges of $G_{n,m}$ until it becomes so simple that we can take a random colouring in polynomial time. Then, we *rebuild* the graph by adding the deleted edges one by one while at the same time we *update the colouring* of some vertices so as the colouring of the graph with the added edge to remain random. This algorithm requires at least $(2 + \epsilon)d$ colours. However, since its accuracy depends on non-reconstruction conditions it is reasonable to expect that we can have an improvement by requiring even less colours. The algorithm does not exploit fully its dependency on non-reconstruction due to its colouring *update rule*. A new, improved, update rule is needed. Such an improvement could possibly reduce the minimum number of colours that the algorithm requires down to $(1 + \epsilon)d / \ln d$.

Very good candidates for improved updating rules are couplings as the one we present here. Of course so as to use it for sampling it requires solving non-trivial technical issues (which go beyond [10, 14]). The close relation between random colourings of d -regular trees and the random colourings of $G_{n,m}$ with expected degree d suggests that such an extension is a reasonable thing to ask. The situation is very similar if someone considers sampling colourings of random d -regular graphs.

Remark 3 Of course, there are other approaches for sampling colouring that vary from combinatorial ones (e.g. Markov Chain Monte Carlo [6, 18]) to numerical ones (e.g. Belief propagation [7]). However, it seems that the setting in [9] is substantially easier to analyse for k being a small function of d .

1.3 A basic description of the coupling.

Assume that we have the d -ary tree T and two broadcasting processes. The first one k -colours T as X and the other as Y . Assume that the root r of T is coloured as $X(r) = c$ and $Y(r) = q$ while $c \neq q$.

Consider first the following recursive *naive coupling* of the two processes. Start from the root r down to the leaves. For each vertex $u \in T$ such that $X(u) = Y(u)$ couple every w , descendant of u ,

¹With a reasonable accuracy

by using identical coupling, i.e. $X(w) = Y(w)$. On the other hand, if $X(u) \neq Y(u)$, then couple the colourings of the children of u according to an optimal coupling for each child.

The above coupling has the property that for any vertex w such that $X(w) \neq Y(w)$, under both X and Y , there is a path which includes r and w and its vertices are coloured using only $X(r)$ and $Y(r)$. That is the *paths of disagreement* in both colourings propagates as bicoloured paths that use the colours c, q . For a number of colours $k < d$, we expect that the above coupling generates an ever increasing number of disagreeing vertices as it moves from the root down to the leaves. As a matter of fact the number of disagreeing vertices at each level grows as a *supercritical* branching process, i.e. the probability of having disagreement at the leaves is strictly positive, regardless of their distance from the root.

Our coupling improves on the naive one. Consider the root r and let N_i be the set of vertices which contain the i -th child of r and the children of i (i.e. two levels below the root). We call “bad” the colouring $X(N_i)$ if $X(i) = q$ and i has a child j such that $X(j) = c$. Similarly, $Y(N_i)$ is bad if $Y(i) = c$ and i has a child j such that $Y(j) = q$. It turns out that the challenge is to deal with these bad colourings. The problem reveals by just considering the naive coupling. There, if $X(N_i)$ is bad, then $Y(N_i)$ should be bad too. For such a pair the identity coupling is precluded and the creation of disagreements is inevitable. For $k < d$, the naive coupling fails due to the fact that it generates *too many* bad pairs $X(N_i), Y(N_i)$.

It is direct to see that for $k < d$ it is unavoidable that there are a lot of bad colouring among $X(N_i)$ s and $Y(N_i)$ s. The basic idea, here, is to avoid coupling them together. To this end we use the following fact which is true as long as $k \geq (2 + \epsilon)d / \ln d$: Consider $X(N_j)$ conditional that A) it is a bad and B) there is at least one colour that is not used by $X(N_j)$. For such $X(N_j)$ it is highly likely that there is another child of r , e.g. vertex s , such that the joint distribution of the colouring of the children of s under $Y(N_s)$ is identical to the joint distribution of the children of j under the colouring $X(N_j)$. For convenience, we call $Y(N_s)$ j -good, since $X(N_j)$ is bad².

Based on the above observation, the target now is to couple the colourings $X(N_i)$ s and $Y(N_i)$ s such that if $X(N_i)$ satisfies the conditions A) and B), stated above, then $Y(N_i)$ is a i -good colouring and vice versa³. Then we can have identical coupling for the colourings of the children of the vertex i . On the other hand, for $X(N_i)$ (or $Y(N_i)$) a bad colouring that either uses all of the k colours or there is not any i -good colouring, it is inevitable that disagreements are generated. For all the rest colourings, i.e. neither bad nor good we have identical coupling. In the next section we provide further details about our coupling on a more technical basis.

Working as described in the previous paragraph, the number of disagreements drops dramatically, compared to naive coupling. As a matter of fact the number of disagreeing vertices grows as a *subcritical* branching process, i.e. the probability of having disagreement at the leaves drops exponentially with the distance of the leaves from the root.

Remark 4 One could hope that employing directly the ideas of this improved coupling to larger neighbourhoods could reduce the bound for k down to $(1 + \epsilon)d / \ln d$. However, technical reasons suggest that a non-trivial improvement is required for getting k below $(2 + \epsilon)d / \ln d$, rather than only considering larger neighbourhoods.

Remark 5 The update rule in the sampling algorithm in [9] is, somehow, an extension to $G_{n,m}$ of what we call here as *naive coupling*. In Section 1.2 we suggest extending this new coupling as an update rule for the algorithm in [9]. This would yield to a lower bound for k below the expected degree d .

²Of course, for the bad colourings among $Y(N_i)$ s we can find good colourings among $X(N_i)$ s, as well.

³If $Y(N_j)$ satisfies the conditions A) and B), then $X(N_j)$ is a j -good colouring.

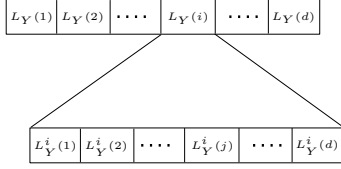


Figure 1: The lists L_Y and L_Y^i .

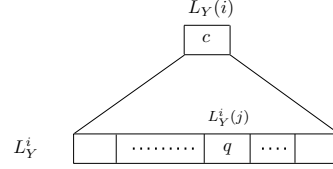


Figure 2: The “Bad” list L_Y^i .

Notation. We use small letters of the greek alphabet for the colourings of G , e.g. σ, τ . The capital letters denote random variables which take values over the colourings e.g. X, Y . We let σ_v denote the colour assignment of the vertex v under the colouring σ . Similarly, the random variable $X(v)$ is equal to the colour assignment that X specifies for the vertex v . For an integer $k > 0$ we let $[k] = \{1, \dots, k\}$.

2 Coupling

In this section we present the coupling in full detail. We consider a complete d -ary tree T . Let $\mu(\cdot)$ denote the uniform distribution over the k -colourings of T , for some k we define later. We consider two broadcasting process such that the first one assigns the root r colour c while the second one assigns the root q . To avoid trivialities assume that $c \neq q$. Finally, we let X, Y be the colourings that the first and the second processes assign to T , respectively. It holds that X is distributed as in $\mu(\cdot | X(r) = c)$ while Y is distributed as in $\mu(\cdot | Y(r) = q)$. We proceed by introducing some useful concepts.

2.1 Preliminaries

Lists of Colour Assignments. We let $L_X \in [k]^d$ be an ordered list which contains the colours that are assigned to the children of the root r under the colouring X , but without revealing which colour gets to which child. Since X is distributed according to $\mu(\cdot | X(r) = c)$ each entry $L_X(i)$ is a random choice from $[k] \setminus \{c\}$. Additionally, for every $i \in [d]$ we let L_X^i be the corresponding lists of the colour assignments of the children of the vertex that is going to be assigned the colour $L_X(i)$. We, also, have the corresponding lists w.r.t the colouring Y (see e.g. Figure 1).

Remark 6 Given L_X , the colouring of the children of r under the colouring X corresponds to a random permutation of the elements of L_X . That is, if π is a random permutation of the elements $(1, \dots, d)$, then for the i -th child of r we can set $X(i) = L_X(\pi(i))$. We work similarly for the grandchildren of r with the lists L_X^i .

Each of L_X^i, L_Y^i will be classified to at least one (possibly more) of the following 4 categories of lists.

bad: We call “bad” every list $L_X^i (L_Y^i)$ which has the property that $c \in L_X^i (q \in L_Y^i)$ while $L_X(i) = q (L_Y(i) = c)$. E.g. see Figure 2.

rescuable: A list L_X^i (or L_Y^i) is rescuable iff it is “bad” and it contains less than $k - 1$ different colours.

j -special: Given a rescuable list L_Y^j , a list L_X^i is called “ j -special” if the following holds:

1. $L_X(i) \neq q$ and $L_X(i) \notin L_Y^j$.
2. One of the following two happens:
 - L_X^i contains q but it does not contain c .
 - L_X^i contains c but it does not contain q .

W.r.t. to a rescuable list L_X^j , we define the j -special lists L_Y^i analogously (e.g. see Figure 3).

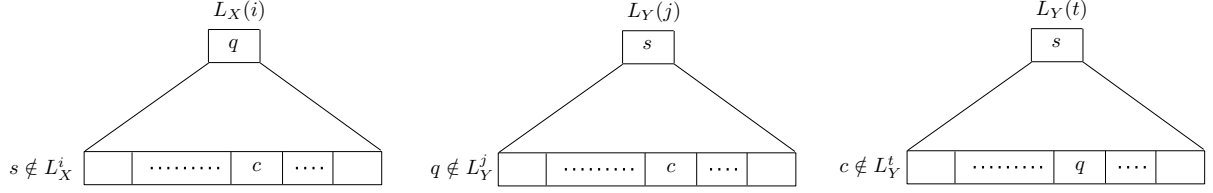


Figure 3: The lists L_Y^j and L_Y^t are both i -special (L_X^i is rescuable). Additionally, the list L_Y^j is i -good.

j -good: A list L_X^i is called “ j -good” if it is j -special and $q \in L_X^i$ while $c \notin L_X^i$. Similarly the j -special L_Y^i is j -good if $c \in L_Y^i$ and $q \notin L_Y^i$ (e.g. see Figure 3).

Lemma 1 Assume that for some $i, j \in [d]$ it holds that L_X^j is rescuable and L_Y^i is j -good. Then L_X^j and L_Y^i are identically distributed.

For the proof of Lemma 1 see in Section 5.1.

2.2 The coupling

For both colourings X, Y consider the lists of colour assignments L_X, L_Y as well as the lists L_X^i and L_Y^i , for $i = 1, \dots, d$. First, we are going to couple the lists. Then, given the colours in the lists we are going to define the actual colouring of the vertices (see Remark 6).

Roughly speaking, the coupling works as follows. In a *first phase* it reveals only a certain part of information about the lists L_X, L_Y, L_X^j and L_Y^j . That is, it reveals the “bad” lists, the “rescuable” and the “special” lists. In a *second phase*, it uses this information about the lists to construct a bijection $f : [d] \rightarrow [d]$ among L_X^i s and L_Y^i s. The use of this bijection is the following: If $f(i) = j$, then when we reveal the full information about the lists we couple optimally $L_X(i)$ with $L_Y(j)$ and L_X^i with L_Y^j , conditional, of course, the information we have already revealed. The bijection f is constructed so as to minimize the number of disagreements at the grandchildren of r . In a *third phase*, we reveal the full information about the lists by using f , as described above. Finally, given the full information for the lists we reveal the assignments of X, Y for the (grand)children of r . Note that if $f(i) = j$ then the child of r that gets $L_X(i)$ under X will get $L_Y(j)$ under Y . Additionally, the grand child of r that is assigned the colour $L_X^i(t)$ under X is going to take the colour $L_Y^j(t)$ under Y .

The pseudocode for the Phase 1 follows:

Phase 1: /*Partial revelation of the lists*/

1. Reveal only for which i we have $L_X(i), L_Y(i) \in \{c, q\}$. Couple the choices of L_X and L_Y such that if $L_X(i) = q$, then $L_Y(i) = c$.
2. For each i such that $L_X(i) = q$ and $L_Y(i) = c$ reveal whether L_X^i and L_Y^i is “bad” or not. We use coupling such that the result of revelation to be the same for both L_X^i and L_Y^i .
3. For each pair of bad lists (L_X^i, L_Y^i) reveal whether they are “rescuable”. We use coupling such that the result of revelation to be the same for both L_X^i and L_Y^i . The coupling is so that *the set of colours in $[k] \setminus \{c, q\}$ that are not used to be the same for both rescuable lists L_X^i, L_Y^i .*
4. Partition the non-bad L_X^j s and L_Y^j s to (maximal) equally sized sets. The partitions are as many as the rescuable pairs and each rescuable pair is associated to exactly one partition. That is, the rescuable pair (L_X^i, L_Y^i) is associated to a set of indices $A_i \subseteq [d]$ such that the following holds: For any $t \in A_i$ $L_X^i(t)$ and $L_Y^i(t)$ belong to partition associated to (L_X^i, L_Y^i) .

5. For each i and for each $j \in A_i$ we reveal if the pair (L_X^j, L_Y^j) consists of i -special lists. We use coupling such that either both lists in the pair are i -special or both are not.

We should recognize the bad lists as sources of potential disagreements in the coupling. However, our attempt is to eliminate the disagreements caused only by the rescuable ones (a subset of bad ones⁴). For this reason each rescuable pair of lists (L_X^i, L_Y^i) is associated with the i -special lists in A_i . Before proceeding, let us make the following clarification.

Remark 7 For the i -special pair (L_X^j, L_Y^j) , above, we do not necessarily have $L_X(j) = L_Y(j)$.

In the next phase we construct the bijection f . Basically, what we need to do is for each rescuable pair (L_X^i, L_Y^i) to find an i -good pair (L_X^j, L_Y^j) among the i -special pairs in A_i . Once we find one, we set $f(i) = j$ and $f(j) = i$. That is, when we reveal the full information of the lists we couple L_X^i with L_Y^j and L_X^j with L_Y^i . Then, from Lemma 1 we can have identity coupling for both pairs. Thus, we eliminate the disagreement we would have had if we had coupled L_X^i and L_Y^i , together.

However, so as to have an i -special pair (L_X^j, L_Y^j) , for $j \in A_i$, with both lists i -good we should couple the lists such that $L_X^j \neq L_Y^j$. To be more specific, we have to reveal whether $c \in L_X^j$ and $q \notin L_X^j$ or the opposite, i.e. $c \notin L_X^j$ and $q \in L_X^j$. Once we have this information for L_X^j , e.g. assume that we have $c \in L_X^j$ and $q \notin L_X^j$, then the coupling should decide the opposite for L_Y^j , i.e. $c \notin L_Y^j$ and $q \in L_Y^j$. It is straightforward that when we reveal an i -special pair (L_X^j, L_Y^j) in such a way it is i -good with probability $1/2$. With the remaining probability we end up with a pair (L_X^j, L_Y^j) which is not i -good, i.e. useless, but even worse we cannot couple L_X^j with L_Y^j identically. Call such a disagreeing pair as i -fail (see example in Figure 4, the upper pair is i -fail). It is straightforward, now, that as we search for an i -good pair it is possible that we generate extra sources of (potential) disagreements. To this end we use the following lemma.

Lemma 2 Assume that the i -special pairs (L_X^t, L_Y^t) and (L_X^s, L_Y^s) in A_i are revealed and (L_X^t, L_Y^t) is i -good while (L_X^s, L_Y^s) is i -fail. Then, L_X^t is identically distributed to L_Y^s and L_Y^t is identically distributed to L_X^s .

For a proof of Lemma 2 see in Section 5.2. Figure 4 gives a schematic representation of what is stated in Lemma 2. The arrows (Figure 4) show the pairs of lists that are identically distributed.

To deal with the potential disagreements generated by i -fails we reveal extra i -good pairs. In particular, we work as follows:

Phase 2. /*List Association*/

For each rescuable pair (L_X^i, L_Y^i) do the following:

1. Reveal each i -special pairs in A_i whether it is i -good or i -fail until either of the following two happens:
 - the number of i -good pairs exceeds the number of i -fails by one,
 - there are no other i -special pairs in A_i to reveal.
2. Reveal the remaining unrevealed i -special pairs, if any, by using identity coupling.
3. Match every i -good pair with one i -fail such that no two i -good pairs are matched to the same i -fail pair.
4. If the i -good pair (L_X^j, L_Y^j) is matched with the i -fail (L_X^s, L_Y^s) , then set $f(j) = s$ and $f(s) = j$.

⁴For the values of k we are interested in, it is highly unlikely that a bad list is non-rescuable.

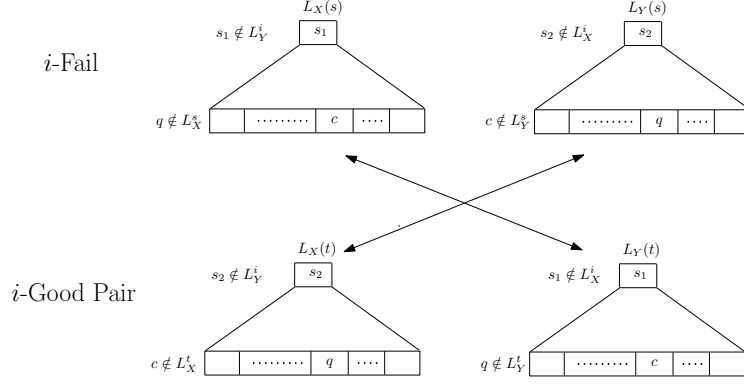


Figure 4: Matching between i -fail and i -good pairs. The coupling between $L_X(s)$ - $L_Y(t)$ and $L_X(t)$ - $L_Y(s)$ is done after the good/fail revelations.

5. If there is an i -good pair (L_X^j, L_Y^j) for which there is no i -fail to be matched, match it with the rescuable pair (L_X^i, L_Y^i) . Also, set $f(i) = j$ and $f(j) = i$.
6. For each $j \in A_i$ such that (L_X^j, L_Y^j) is not matched yet, match it to itself and set $f(j) = j$.

Ideally, Phase 2 generates a number of i -good pairs which exceed the number of i -fails by one. If this is the case, then from Lemma 2 we can construct the function f such that no disagreement is generated by any i -fail pair in the full revelation of the lists. Furthermore, there is an extra i -good pair to be matched with the rescuable pair (L_X^i, L_Y^i) and no disagreement is generated by the rescuable pair as well (by Lemma 1).

Of course, it is possible that the number of the i -good pairs is not sufficiently large. Then, we end up with the rescuable pair (L_X^i, L_Y^i) with some i -fails which cannot be matched with any i -good pair. These pairs are matched to themselves and some disagreements are going to appear in the full revelation. However, we show that the expected number of disagreements vanishes as long as $k \geq (2 + \epsilon)d / \ln d$. We now, proceed with the Phase 3.

Phase 3: /*Full revelation*/

1. For every s, t such that $f(s) = t$, couple optimally $L_X(s)$ with $L_Y(t)$ as well as L_X^s with L_Y^t .
2. Reveal which element of the list L_X is assigned to which child of r and which element of L_X^j goes to which grandchild of r , as Remark 6 specifies.
3. Assuming that v , child of r , is such that $X(v)$ is set $L_X(s)$, then we set $Y(v)$ equal to $L_Y(t)$, where $t = f(s)$. Also, for u , child of v , such that $X(u)$ set $L_X^s(j)$ we set $Y(u)$ equal to $L_Y^t(j)$.

Applying the same coupling for the grandchildren of the root, at the end we get the full colourings X and Y . For completeness, in Section 6 we provide the full pseudocode of coupling as a recursive procedure. A very basic result is the following theorem.

Theorem 2 For $c, q \in [k]$, assume that in the above coupling it holds $X(r) = c$ and $Y(r) = q$, where r is the root vertex of T . Then at the end of the coupling, X and Y are distributed as in $\mu(\cdot | X(r) = c)$ and $\mu(\cdot | Y(r) = q)$, respectively.

Proof: Theorem follows by noting that for every list, conditional on the information that is already known to us, we reveal some information by using the appropriate distribution. \square

Furthermore, from the description of the coupling the following corollary is direct.

Corollary 2 *The disagreements in the coupling have three different sources:*

1. *Pairs of bad lists which are not rescuable.*
2. *Pairs of rescuable lists for which it was impossible to find a good pair.*
3. *Pairs of i -fail lists, for some i , which are not matched to an i -good pair.*

Proposition 1 *Consider the above coupling between X and Y and let \mathcal{W}_l be the number of vertices u at level l such that $X(u) \neq Y(u)$. For fixed $\epsilon > 0$, sufficiently large d , $k = (1 + \epsilon)\frac{d}{\ln d}$ and every even integer $l > 0$ it holds that*

$$E[\mathcal{W}_l] \leq \left(d^{-0.1 \frac{\epsilon-1}{\epsilon+1}} \right)^{l/2}.$$

Using Proposition 1 it is direct to see that our combinatorial construction implies the following theorem.

Theorem 3 *For fixed $\epsilon > 0$ and sufficiently large d , the following is true for the k -colouring model on a d -ary tree T : If $k = (2 + \epsilon)d/\ln d$, then the model is non-reconstructible.*

Proof: Take $k = (2 + \epsilon)d/\ln d$. Let X and Y be distributed as in $\mu(\cdot|X(r) = c)$ and $\mu(\cdot|Y(r) = q)$, respectively, while their joint distribution is specified by the coupling we presented. Let L_h be the set of vertices that contains all the vertices of T at distance h . We take h even. By Coupling Lemma we have

$$\|\mu(\cdot|X(r) = c) - \mu(\cdot|Y(r) = q)\|_{L_h} \leq \Pr[X(L_h) \neq Y(L_h)]. \quad (1)$$

Let \mathcal{W}_h be the number of vertices $u \in L_h$ such that $X(u) \neq Y(u)$. It holds that

$$\begin{aligned} \Pr[X(L_h) \neq Y(L_h)] &= \Pr[\mathcal{W}_h > 0] \leq E[\mathcal{W}_h] && \text{[by Markov's inequality]} \\ &\leq \left(d^{-0.1 \frac{\epsilon}{\epsilon+2}} \right)^{\lfloor h/2 \rfloor} && \text{[from Proposition 1].} \end{aligned} \quad (2)$$

The theorem follows by combining (1) and (2). \square

3 Proof of Proposition 1

Consider in the coupling two vertices $v, w \in T$ at the same level l , where l is even. Given the colourings $X(v)$, $Y(v)$ and $X(w)$ and $Y(w)$, and assuming that $X(v) \neq Y(v)$ and $X(w) \neq Y(w)$ observe that whether the descendants of v disagree or not does not depend on what happens at the descendants of w and vice versa. This observation yields to the following: For some vertex $v \in T$ that $X(v) \neq Y(v)$, let \mathcal{D}_v be the number of disagreements two levels below the vertex v . Then, it holds that

$$E[\mathcal{W}_l | \mathcal{W}_{l-2}] = E[\mathcal{D}_v] \cdot \mathcal{W}_{l-2} \quad \text{for even } l > 0.$$

Taking the average from both sides and working out the recursion it is easy to get that

$$E[\mathcal{W}_l] = (E[\mathcal{D}_v])^{l/2}.$$

The proposition will follow by bounding appropriately $E[D_v]$. So as to bound $E[D_v]$ we need to bound the number of disagreements that are generated by each of the three sources of disagreement specified in Corollary 2. It, always, holds that $D_v \leq d^2$, since T is a d -ary tree.

Consider the following quantities related to the vertex v : Let β_v denote the number of bad lists. Let δ_k be the probability for a bad list to be rescuable, for a given number of colours k . Finally, given some rescuable list L_X^j let h_v^j be the number of j -special lists in the partition it is assigned. Let the event \mathbb{A} denote that at least one of the following three is true

1. $\beta_v \geq 100 \ln d$.
2. There is at least one bad list which is not rescuable.
3. There is a rescuable list L_X^j that is corresponded to a partition with less than $d^{\frac{4}{5} \frac{\epsilon-1}{1+\epsilon}}$ j -special lists.

It is direct to get that

$$E[D_v] \leq d^2 Pr[\mathbb{A}] + E[D_v | \mathbb{A}^c], \quad (3)$$

where we use the rather crude overestimate that conditional on the event \mathbb{A} occurs all the d^2 descendants of v are disagreeing. It suffices to bound appropriately $Pr[\mathbb{A}]$ and $E[D_v | \mathbb{A}^c]$. To this end, we use the following propositions.

Proposition 2 For $k = (1 + \epsilon)d / \ln d$ and for sufficiently large d , we have that

$$E[D_v | \mathbb{A}^c] \leq d^{-0.102 \frac{\epsilon-1}{\epsilon+1}}.$$

Proposition 3 For $k = (1 + \epsilon)d / \ln d$ and for sufficiently large d , we have that

$$Pr[\mathbb{A}] \leq 5d^{-250}.$$

Form Proposition 2, Proposition 3 and for sufficiently large d , we have that (3) implies that

$$E[D_v] \leq d^{-0.1 \frac{\epsilon-1}{\epsilon+1}}.$$

The proposition follows.

3.1 Proof of Proposition 2

Since we have conditioned on \mathbb{A}^c , we have that A) β_v , the number of bad lists, is at most $100 \ln d$, B) all the bad lists are rescuable and C) every rescuable list L_X^i is assigned to A_i which contains at least $d^{\frac{4}{5} \frac{\epsilon-1}{1+\epsilon}}$ i -special lists. Clearly, from (A) and (B) we deduce that the number of rescuable lists is equal to β_v .

Let the set Δ_i contain the pair (L_X^s, L_Y^s) , for $s \in A_i \cup \{i\}$, if L_X^s should be coupled with L_Y^s in Phase 3 while identical coupling of the lists is impossible. We remind the reader that in the second phase of the coupling we reveal which of the i -special pairs in A_i are i -good or not, i.e. during the steps 1 and 2. During these revelations it is possible that we introduce pairs which are i -fails which are coupled together (due to lack of i -good pair). These i -fails, for which it was imposible to find i -good pair belong to Δ_i . Of course, the rescuable pair (L_X^i, L_Y^i) can also be in Δ_i , so long as L_X^i should be coupled with L_Y^i in Phase 3.

Conditional on the event \mathbb{A}^c the following holds: Consider $(L_X^t, L_Y^t) \in \Delta_i$ which is non-rescuable. It is easy to see that we can couple $L_X(t)$, $L_Y(t)$ such that $L_X(t) = L_Y(t)$. Also, given that $L_X(t) = L_Y(t)$ it holds that $c \in L_X^t$ and $q \notin L_X^t$, also $q \in L_Y^t$ and $c \notin L_Y^t$. Furthermore, all the rest colours, i.e. all but $c, q, L_X(t)$, are symmetric for both lists L_X^t and L_Y^t . Clearly, we can couple L_X^t and L_Y^t , such that if $L_X^t(s) = c$ then $L_Y^t(s) = q$ while if $L_X^t(s) \neq c$, then $L_X^t(s) = L_Y^t(s)$ for any $s \in [d]$. Also, we can have exactly the same coupling for the rescuable pair $(L_X^i, L_Y^i) \in \Delta_i$.

Let Z_t be the number of disagreements that are generated by the coupling of the pair $(L_X^t, L_Y^t) \in \Delta_i$. Also, let Q_i be the number of all disagreements that we get from coupling the pairs in Δ_i after the third phase. It holds that

$$E[Q_i | \mathbb{A}^c] = E[|\Delta_i| | \mathbb{A}^c] \cdot E[Z_j | \mathbb{A}^c].$$

Also, by the linearity of expectation it holds that

$$\begin{aligned} E[D_v | \mathbb{A}^c] &\leq (100 \ln d) E[Q_i | \mathbb{A}^c] && [\text{as } \mathbb{A}^c \text{ assumes that } \beta_v \leq 100 \ln d] \\ &\leq (100 \ln d) \cdot E[|\Delta_i| | \mathbb{A}^c] \cdot E[Z_j | \mathbb{A}^c]. \end{aligned} \quad (4)$$

The proposition will follow by bounding appropriately $E[|\Delta_i| | \mathbb{A}^c]$ and $E[Z_j | \mathbb{A}^c]$.

As far as $E[Z_j | \mathbb{A}^c]$ is concerned notice the following: Conditional on \mathbb{A}^c , every list $L_X^t \in \Delta_i$ has a number of entries with colour c that is binomial distributed $\mathcal{B}(d, 1/k)$, conditional that there exists at least one entry which is c . For $L_Y^t \in \Delta_i$ we have the same conditions w.r.t. the colour q . Coupling the lists $(L_X^t, L_Y^t) \in \Delta_i$ as specified above, the number of disagreements is exactly the number of occurrences of c in L_X^t (or the occurrences of q in L_Y^t). Then, it follows easily that

$$\begin{aligned} E[Z_j | \mathbb{A}^c] &= \sum_{s=0}^d s \cdot \Pr[c \text{ appears } s \text{ times in } L_X^t | c \text{ appears at least once in } L_X^t] \\ &= \frac{1}{1 - \left(1 - \frac{1}{k-1}\right)^d} \sum_{s=1}^d s \cdot \binom{d}{s} \left(\frac{1}{k-1}\right)^s \left(1 - \frac{1}{k-1}\right)^{d-s} \\ &\leq \frac{d}{k-1} && [\text{since } 1 - \left(1 - \frac{1}{k-1}\right)^d > 1/2] \\ &\leq 2 \ln d. && [\text{since } k = (1 + \epsilon)d / \ln d] \end{aligned} \quad (5)$$

As far as $E[|\Delta_i| | \mathbb{A}^c]$ is concerned, we work as follows: Let S_i be the set that contains all the i -special pairs in A_i as well as the rescuable pair (L_X^i, L_Y^i) . W.l.o.g. assume that $i = 1$ while the indices of the i -special pairs in S_i are from 2 to $|S_i|$. Let the 0-1 matrix $\mathcal{S} = |S_i| \times 2$ be defined as follows: $\mathcal{S}(1, t) = 1$, if $c \in L_X^t$ and $q \notin L_X^t$, otherwise, i.e. $c \notin L_X^t$ and $q \in L_X^t$, $\mathcal{S}(1, t) = 0$. Similarly, $\mathcal{S}(2, t) = 1$ if $c \notin L_Y^t$ and $q \in L_Y^t$, otherwise $\mathcal{S}(2, t) = 0$. It is clear that for the i -special pair (L_X^t, L_Y^t) that is i -good, it holds that $(\mathcal{S}(1, t), \mathcal{S}(2, t)) = (0, 1)$ on the other hand if (L_X^t, L_Y^t) is i -fail, then $(\mathcal{S}(1, t), \mathcal{S}(2, t)) = (1, 0)$.

Remark 8 The second phase of the coupling specifies how $\mathcal{S}(1, j)$ and $\mathcal{S}(2, j)$ are correlated with each other. In particular, the following holds: if $\sum_{j=1}^{i-1} (\mathcal{S}(1, j) - \mathcal{S}(2, j)) > 0$, then $\mathcal{S}(1, i)$ and $\mathcal{S}(2, i)$ are coupled so as to get complementary values. Otherwise, i.e. if $\sum_{j=1}^{i-1} (\mathcal{S}(1, j) - \mathcal{S}(2, j)) = 0$, they are coupled identically.

Since we have assumed that the values in $(\mathcal{S}(1, 1), (2, 1))$ are related to (L_X^i, L_Y^i) , by definition it holds that $(\mathcal{S}(1, 1), (2, 1)) = (1, 0)$. Furthermore, for each $t = 2 \dots |S_i|$ and as long as $\sum_{j=1}^{t-1} (\mathcal{S}(1, j) - \mathcal{S}(2, j)) > 0$ we have

$$(\mathcal{S}(1, t), \mathcal{S}(2, t)) = \begin{cases} (1, 0) & \text{with probability } 1/2 \\ (0, 1) & \text{with probability } 1/2. \end{cases}$$

For the matrix \mathcal{S} we have the following lemma.

Lemma 3 The cardinality of Δ_i and \mathcal{S} are related as follows:

$$|\Delta_i| = \sum_{t=1}^N \mathcal{S}(1, t) - \mathcal{S}(2, t),$$

where N is the number of columns of the matrix \mathcal{S} .

Proof: First notice that $S(1, 1) - S(2, 1) = 1$. The coupling during the second phase assigns complementary values to each pair $S(1, t), S(2, t)$ as long as $R_t = \sum_{i=1}^{t-1} [S(1, i) - S(2, i)] > 0$. Once $R_t = 0$ it sets $S(1, t) = S(2, t)$, i.e. R_t remains zero for the rest values of t .

Let T be the maximum t such that $S(1, t) \neq S(2, t)$. Clearly, it suffice to show that

$$|\Delta_i| = \sum_{t=1}^T \mathcal{S}(1, t) - \mathcal{S}(2, t).$$

For $t < T$, the fact that $\mathcal{S}(1, t) = 1$ (and consequently $\mathcal{S}(2, t) = 0$) suggests that we have revealed an i -fail. On the other hand, if $\mathcal{S}(1, t) = 0$ (and consequently $\mathcal{S}(2, t) = 1$), then it suggests that it has been revealed an i -good pair. This observation implies that the sum $\sum_{t=1}^T \mathcal{S}(1, t)$ is equal to the number of i -fails we have revealed, while $\sum_{t=1}^T \mathcal{S}(2, t)$ is equal to the number of i -good pairs.

Since we can match an i -fail with an i -good pair to avoid generating disagreements, the number of pairs which do not admit identical coupling, i.e. the i -fail and possibly the rescuable pair, is equal to

$$\sum_{t=1}^T \mathcal{S}(1, t) - \mathcal{S}(2, t) = \sum_{t=1}^N \mathcal{S}(1, t) - \mathcal{S}(2, t).$$

The lemma follows. \square

Proposition 4 *Let N be the number of columns of \mathcal{S} . Then for sufficiently large N it holds that*

$$E \left[\sum_{j=1}^N (\mathcal{S}(1, j) - \mathcal{S}(2, j)) \right] \leq \left(\frac{2.3}{\pi} \right)^{0.43 \ln N}. \quad (6)$$

For a proof of Proposition 4 see in Section 4. Using Proposition 4 and Lemma 3 and the assumption that the number of i -special pairs in A_i is at least $d^{\frac{4}{5} \frac{\epsilon-1}{\epsilon+1}}$, we get

$$E[|\Delta_i| | \mathbb{A}^c] \leq \left(\frac{2.3}{\pi} \right)^{0.43 \frac{4(\epsilon-1)}{5(\epsilon+1)} \ln d} \leq d^{-0.344 \frac{\epsilon-1}{\epsilon+1} \ln(\frac{\pi}{2.3})} \leq d^{-0.107 \frac{\epsilon-1}{\epsilon+1}}. \quad (7)$$

Plugging into (4) the inequalities (5) and (7) we get that

$$E[D_v | \mathbb{A}^c] \leq \frac{200 \ln^2 d}{d^{0.107 \frac{\epsilon-1}{\epsilon+1}}}.$$

The proposition follows by taking sufficiently large d .

3.2 Proof of Proposition 3

For the quantities, β_v, δ_k and h_v we defined in Section 3 we have the following proposition.

Proposition 5 *For $k = (1 + \epsilon)d / \ln d$ the following are true:*

$$Pr \left[\beta_v \geq (1 + x) \frac{d}{k - 1} \right] \leq d^{\left(-\frac{3\phi(x)}{4(1+\epsilon)} \right)}, \quad (8)$$

where $\phi(x) = (1 + x) \ln(1 + x) - x$, for real $x > 0$. Also, it holds that

$$\delta_k \geq 1 - \exp \left(-\frac{3(1 + \epsilon)}{8 \ln d} d^{\frac{\epsilon}{1+\epsilon}} \right). \quad (9)$$

Finally, for any $c > 0$ it holds that

$$Pr \left[h_v \leq \frac{d^{\frac{\epsilon-1}{1+\epsilon}}}{2c \ln d} | \beta_v \leq c \ln d \right] \leq \exp \left(-\frac{d^{\frac{\epsilon-1}{1+\epsilon}}}{8c \ln d} \right). \quad (10)$$

Let the events $E_1 = \beta_v \geq 100 \ln d$, $E_2 = \text{“there is at least one bad list which is not rescuable”}$ and $E_3 = \text{“there is a rescuable list } L_X^j \text{ that is corresponded to a partition with less than } d^{\frac{4}{5} \frac{\epsilon-1}{1+\epsilon}} \text{ } j\text{-special lists”}$. It is direct that

$$Pr[\mathbb{A}] = Pr \left[\bigcup_{i=1}^3 E_i \right] \leq \sum_{i=1}^3 Pr[E_i]. \quad (11)$$

The proposition will follow by bounding appropriately the probability terms $Pr[E_1]$, $Pr[E_2]$ and $Pr[E_3]$ in the above relation. As far as $Pr[E_1]$ is regarded note that for $1 + x_0 = 98(1 + \epsilon)$ it holds that

$$Pr[E_1] \leq Pr \left[\beta_v > (1 + x_0) \frac{d}{k-1} \right]. \quad (12)$$

The above inequality holds since

$$\begin{aligned} \frac{d}{k-1} &\leq \frac{d}{k} + \frac{2d}{k^2} && [\text{as } \frac{1}{k-1} \leq \frac{1}{k} + \frac{2}{k^2}] \\ &\leq \frac{\ln d}{(1 + \epsilon)} + \frac{2 \ln^2 d}{d} && [\text{as } k \geq (1 + \epsilon)d / \ln d]. \end{aligned}$$

We use Proposition 5, (i.e. (8)) to bound the r.h.s of (12). Noting that for $x_0 = 98(1 + \epsilon) - 1$ it holds that $\phi(x_0) \geq 343(1 + \epsilon) + 98(1 + \epsilon) \ln(1 + \epsilon)$, it is direct to see that

$$Pr[E_1] \leq d^{-250}. \quad (13)$$

As far as $Pr[E_2]$ is regarded, we let NR_v be the number of non-rescuable lists. Clearly, it holds that

$$Pr[E_2] = Pr[NR_v > 0] \leq E[NR_v], \quad (14)$$

where the last inequality follows from Markov's inequality. Using (9), it is direct that

$$\begin{aligned} E[NR_v] &\leq (1 - \delta_k) d \\ &\leq \exp \left(-\frac{3(1 + \epsilon)}{8 \ln d} d^{\frac{\epsilon}{1+\epsilon}} \right) d \leq \exp \left(-d^{\frac{\epsilon}{2(1+\epsilon)}} \right). \end{aligned}$$

Plugging the above inequality to (14) we get that

$$Pr[E_2] \leq \exp \left(-d^{\frac{\epsilon}{2(1+\epsilon)}} \right). \quad (15)$$

Finally as far as $Pr[E_3]$ is regarded, note that

$$Pr[E_3] \leq Pr[E_3 | \beta_v < 100 \ln d] + Pr[\beta_v \geq 100 \ln d]. \quad (16)$$

We let NC_v be the number of bad lists which are corresponded to a partition with less than $d^{\frac{4}{5} \frac{\epsilon-1}{1+\epsilon}}$ special lists. Clearly, it holds that

$$Pr[E_3 | \beta_v \leq 100 \ln d] = Pr[NC_v > 0 | \beta_v \leq 100 \ln d].$$

For the rescuable list L^j we let h_v denote the number of j -special lists that is assigned. It holds that

$$\begin{aligned}
Pr[h_v \leq d^{\frac{4}{5}\frac{\epsilon-1}{1+\epsilon}} | \beta_v \leq 100 \ln d] &\leq Pr\left[h_v \leq \frac{d^{\frac{\epsilon-1}{1+\epsilon}}}{4 \ln^5 d} | \beta_v \leq 100 \ln d\right] \\
&\leq \exp\left(-\frac{d^{\frac{\epsilon-1}{1+\epsilon}}}{800 \ln d}\right) \quad [\text{from (10)}] \\
&\leq \exp\left(-d^{\frac{4}{5}\frac{\epsilon-1}{1+\epsilon}}\right).
\end{aligned}$$

It is direct that

$$\begin{aligned}
E[NC_v | \beta_v \leq 100 \ln d] &\leq (100 \ln d) Pr[h_v \leq d^{\frac{4}{5}\frac{\epsilon-1}{1+\epsilon}} | \beta_v \leq 100 \ln d] \\
&\leq \exp\left(-d^{\frac{3}{5}\frac{\epsilon-1}{1+\epsilon}}\right).
\end{aligned}$$

Using Markov's inequality we get that

$$Pr[NC_v > 0 | \beta_v \leq 100 \ln d] \leq E[NC_v | \beta_v \leq 100 \ln d] \leq \exp\left(-d^{\frac{3}{5}\frac{\epsilon-1}{1+\epsilon}}\right).$$

Plugging the above inequality and (13) to (16) we get that

$$Pr[E_3] \leq \exp\left(-d^{\frac{3}{5}\frac{\epsilon-1}{1+\epsilon}}\right) + d^{-250} \leq 2d^{-250}. \quad (17)$$

Plugging in (13), (15) and (17) to (11) we get that $Pr[\mathbb{A}] \leq 5d^{-250}$. The proposition follows.

3.3 Proof of Proposition 5

The inequality in (9) follows from the following two lemmas.

Lemma 4 Consider a random k -colouring of T . Let $v \in T$, and let L be the list of colours that appear in the children of v . For $k = (1 + \epsilon)d / \ln d$ and for any colour c that is not assigned to v it holds that

$$|Pr[c \notin L] - d^{-\frac{1}{1+\epsilon}}| \leq d^{-\frac{1}{1+\epsilon}} \cdot \frac{4d}{k^2}.$$

Proof: Clearly, it holds that

$$\begin{aligned}
Pr[c \notin L] &= \left(1 - \frac{1}{k-1}\right)^d \leq \exp\left(-\frac{d}{k-1}\right) \quad [\text{as } 1 - x \leq e^{-x}] \\
&\leq \exp\left(-\frac{d}{k}\right) \leq d^{-\frac{1}{1+\epsilon}}.
\end{aligned}$$

Also we have that

$$\begin{aligned}
Pr[c \notin L] &= \left(1 - \frac{1}{k-1}\right)^d \geq \exp\left(-\frac{d}{k-2}\right) \quad [\text{as } 1 - x \geq e^{-x/(1-x)} \text{ for } 0 < x < 0.1] \\
&\geq \exp\left(-\frac{d}{k} - \frac{4d}{k^2}\right) \geq d^{-\frac{1}{1+\epsilon}} \left(1 - \frac{4d}{k^2}\right) \quad \left[\text{as } \frac{1}{k-2} \leq \frac{1}{k} + \frac{4}{k^2}\right].
\end{aligned}$$

□

Lemma 5 For a vertex v let f_v denote the number of colours that are not used by its children under a random colouring. For $k = (1 + \epsilon)d / \ln d$ and for any $y \in (0, 1)$ it holds that

$$Pr \left[f_v \leq (1 - y) \frac{3(1 + \epsilon)}{4 \ln d} d^{\frac{\epsilon}{1+\epsilon}} \right] \leq \exp \left(- \frac{3y^2 (1 + \epsilon)}{8 \ln d} d^{\frac{\epsilon}{1+\epsilon}} \right). \quad (18)$$

Proof: Since v is assigned some colour, the number of available colours for its children is $k - 1$. Using Lemma 4 and linearity of expectation we get that

$$\begin{aligned} E[f_v] &\geq (k - 1) d^{-\frac{1}{1+\epsilon}} (1 - 4d/k^2) \\ &\geq (1 + \epsilon) \frac{d^{\frac{\epsilon}{1+\epsilon}}}{\ln d} \left(1 - 8 \frac{\ln d}{d} \right) \geq \frac{3(1 + \epsilon)}{4 \ln d} d^{\frac{\epsilon}{1+\epsilon}}. \end{aligned} \quad (19)$$

Using a ball and bins argument, we can show that we can apply Chernoff bounds for f_v . In particular, for any $y \in (0, 1)$ it holds that

$$Pr[f_v \leq (1 - y)E[f_v]] \leq \exp \left(- \frac{y^2}{2} E[f_v] \right) \leq \exp \left(- \frac{y^2}{2} \frac{3(1 + \epsilon)}{4 \ln d} d^{\frac{\epsilon}{1+\epsilon}} \right) \quad [\text{from (19)}].$$

Clearly, from (19) we get that

$$Pr[f_v \leq (1 - y)E[f_v]] \geq Pr \left[f_v \leq (1 - y) \frac{3(1 + \epsilon)}{4 \ln d} d^{\frac{\epsilon}{1+\epsilon}} \right].$$

Combining the two inequalities above we get (18). \square

It is direct to see that (9) follows from (18), where we set $y = 1/2$, and by noting that

$$\begin{aligned} \delta_k &\geq 1 - Pr \left[f_v \leq \frac{3(1 + \epsilon)}{8 \ln d} d^{\frac{\epsilon}{1+\epsilon}} \right] \\ &\geq 1 - \exp \left(- \frac{3(1 + \epsilon)}{32 \ln d} d^{\frac{\epsilon}{1+\epsilon}} \right), \end{aligned}$$

Also, (8) follows as a corollary from the following lemma.

Lemma 6 Let β_v be the number of bad lists. For $k = (1 + \epsilon)d / \ln d$, it holds that

$$E[\beta_v] \leq \frac{d}{k - 1} \leq \frac{\ln d}{1 + \epsilon} + \frac{\ln^2 d}{d}, \quad (20)$$

$$Pr[\beta_v \geq (1 + x)E[\beta_v]] \leq d^{\left(-\frac{3\phi(x)}{4(1+\epsilon)} \right)} \quad (21)$$

where $\phi(x) = (1 + x) \ln(1 + x) - x$, for $x > 0$.

Proof: Clearly there are d different lists and each of them is bad independently of the others. Let p_{bad} be the probability for the list L^i to be bad. This means that $L_Y(i) = c$ while $q \in L_Y^i$. It holds that

$$p_{bad} = \frac{1}{k - 1} \left(1 - \left(1 - \frac{1}{k - 1} \right)^d \right) \leq \frac{1}{k - 1},$$

as $\left(1 - \left(1 - \frac{1}{k - 1} \right)^d \right) \leq 1$. By linearity of expectation we get that

$$\begin{aligned} E[\beta_v] \leq d p_{bad} &\leq \frac{d}{k - 1} \leq d \left(\frac{1}{k} + \frac{2}{k^2} \right) && \left[\text{as } \frac{1}{k - 1} \leq \frac{1}{k} + \frac{2}{k^2} \right] \\ &\leq \frac{\ln d}{1 + \epsilon} + \frac{\ln^2 d}{d}. \end{aligned}$$

Also, using Lemma 4 we get that

$$p_{bad} \geq \frac{1}{k-1} \left(1 - d^{-\frac{1}{1+\epsilon}} \left(1 + \frac{4d}{k^2} \right) \right) \geq \frac{3}{4k}.$$

In turn, we get that

$$E[\beta_v] \geq dp_{bad} \geq \frac{3 \ln d}{4(1+\epsilon)}. \quad (22)$$

Applying Chernoff bounds we have that for any $x > 0$

$$\Pr[\beta_v \geq (1+x)E[\beta_v]] \leq \exp(-\phi(x) \cdot E[\beta_v]),$$

where $\phi(x) = (1+x) \ln(1+x) - x$. We get (21) by substituting the expectation term on the r.h.s. above by using the bound (22). The lemma follows. \square

The next two lemmas are sufficient to show (10).

Lemma 7 *Let L_X^j be a rescuable list and let A_j be the set of indices where we check for j -special lists. Conditional that A_j is non empty, for some $i \in A_j$, let ϱ_k be the probability for L_Y^i to be j -special. Then for $k \geq (1+\epsilon)d/\ln d$ it holds that*

$$\varrho_k \geq \frac{10}{9} d^{-\frac{2}{1+\epsilon}}$$

Proof: Since L_X^j is rescuable it means that $L_X(j) = c$ and $q \in L_X^j$ while there are colours, apart from c , that do not appear in L_X^j . Let f_j be the number of colours, apart from c , that do not appear in L_X^j , without conditioning that it is rescuable. Furthermore, so as to have L_Y^i j -special, it should hold that $L_Y(i) \notin L_X^j \cup \{c\}$ and either of the following two holds A) $q \in L_Y^i$ and $c \notin L_Y^i$ or B) $q \notin L_Y^i$ and $c \in L_Y^i$. Let the event $\mathbb{Q} = "L_Y(i) \notin L_X^j \cup \{c\}"$. It holds that

$$\varrho_k \geq 2 \frac{E[f_j | f_j > 0]}{k-1} \Pr[q \notin L_Y^j | c \in L_Y^j, \mathbb{Q}] \Pr[c \in L_Y^j | \mathbb{Q}]. \quad (23)$$

Noting that $E[f_j] = E[f_j | f_j > 0] \Pr[f_j > 0]$ we get that

$$E[f_j | f_j > 0] \geq E[f_j] \geq (k-1) d^{-\frac{1}{1+\epsilon}} \left(1 - \frac{4d}{k^2} \right), \quad (24)$$

where the last inequality follows from Lemma 4. Also, by Lemma 4 we get that

$$\Pr[c \in L_Y^i | \mathbb{Q}] \geq 1 - d^{-\frac{1}{1+\epsilon}} \left(1 + \frac{4d}{k^2} \right). \quad (25)$$

Working as in the proof of Lemma 4 we get that

$$\Pr[q \notin L_Y^j | c \in L_Y^j, \mathbb{Q}] \geq \frac{3}{4} d^{-\frac{1}{1+\epsilon}} \quad (26)$$

Substituting the bounds (25), (26) and (24) in (23) we get

$$\varrho_k \geq \frac{3}{2} d^{-\frac{2}{1+\epsilon}} \left(1 - \frac{4d}{k^2} \right)^2 \left(1 - d^{-\frac{1}{1+\epsilon}} \left(1 + \frac{4d}{k^2} \right) \right) \geq \frac{10}{9} d^{-\frac{2}{1+\epsilon}}.$$

\square

Lemma 8 Consider a random k -colouring of T for $k = (1 - \epsilon)d / \ln d$. Let h_v^j be the number of good lists that correspond to each rescuable list L_X^j . For any real $c > 0$, it holds that

$$\Pr \left[h_v \leq \frac{d^{\frac{\epsilon-1}{1+\epsilon}}}{2c \ln d} \mid \beta_v \leq c \ln d \right] \leq \exp \left(-\frac{d^{\frac{\epsilon-1}{1+\epsilon}}}{8c \ln d} \right).$$

Proof: It is direct that the number of lists that are assigned to each rescuable list depends on the actual number of bad lists. Let β_v be the number of bad lists. Conditioning that $\beta_v \leq c \ln d$, for some fixed $c > 0$, it is direct that each rescuable list L_X^j , is assigned a set of at least $\lfloor \frac{d}{c \ln d} - 1 \rfloor$ non-bad lists to find a j -special. Using Lemma 7 we get that

$$E[h_v \mid \beta_v \leq c \ln d] \geq \frac{10}{9} d^{-\frac{2}{1+\epsilon}} \left(\frac{d}{c \ln d} - 2 \right) \geq \frac{d^{\frac{\epsilon-1}{1+\epsilon}}}{c \ln d}.$$

Also, note that given the rescuable lists, each of the remaining lists are special independently of the other lists. Thus, we can apply Chernoff bounds and get the following inequality.

$$\Pr \left[h_v \leq (1 - y) \frac{d^{\frac{\epsilon-1}{1+\epsilon}}}{c \ln d} \mid \beta_v \leq c \ln d \right] \leq \exp \left(-\frac{y^2 d^{\frac{\epsilon-1}{1+\epsilon}}}{2c \ln d} \right).$$

The lemma follows by setting above $y = 1/2$. □

4 Proof of Proposition 4

The matrix \mathcal{S} has random entries, let N be its total number of columns. A way of constructing \mathcal{S} , which is equivalent to the one described in Remark 8, is the following one: Consider some sufficiently large positive integer $l \ll N$. We construct \mathcal{S} in rounds. Assume that after round $i - 1$ we have constructed \mathcal{S} up to some column t , for some $t \ll N$. Additionally, let $X_t = \sum_{j=1}^t \mathcal{S}(1, j) - \mathcal{S}(2, j)$. Then, during the round i we proceed as described in the following paragraph.

If $X_t = 0$, then we use identical coupling for $\mathcal{S}(1, j), \mathcal{S}(2, j)$ for all $t < j \leq N$. If $X_t > 0$, then we consider X_t many sets of columns of \mathcal{S} whose values has not been set yet. Each of these X_t many sets contains at most l columns. More specifically, the first set R_1^i starts from column $t + 1$ up to column T , the value of T will be defined in what follows. We set the values in each column $j \in R_1^i$ by coupling $\mathcal{S}(1, j)$ $\mathcal{S}(2, j)$ such that $\mathcal{S}(1, j) = 1 - \mathcal{S}(2, j)$. T is either the first time that $\sum_{j=t+1}^T \mathcal{S}(1, j) - \mathcal{S}(2, j) = -1$ or if this is not possible up to column $t + l$, then we have $T = t + l$. Continue with the second set of columns R_2^i ⁵ and so on. Round i ends after having finished with all these X_t sets of columns. Then we continue in the same manner with the round $i + 1$.

For each set of columns R_j^i , (R_j^i is submatrix of \mathcal{S}), we have the following lemma which is going to be useful in the proof of Proposition 4.

Lemma 9 Let $l \geq 10$, the maximum number of columns of R_j^i . If the entries are such $R_j^i(1, s) \neq R_j^i(2, s)$ for any column s of R_j^i , then it holds that

$$E \left[1 + \sum_{t=1}^T R_j^i(1, t) - R_j^i(2, t) \right] \leq \frac{2.3}{\pi},$$

where T is the actual number of columns of R_j^i .

⁵ R_2^i starts from the column $T + 1$

Proof: For every t it holds that $R_j^i(1, t) - R_j^i(2, t)$ is equal to -1 with probability $1/2$ or it is equal to 1 with probability $1/2$. It is direct to see that the partial sums $W_s = \sum_{t=1}^s R_j^i(1, t) - R_j^i(2, s)$, for $s \leq T$ constitute a symmetric random walk on the integers which starts from position zero and stops either when it hits -1 or after l steps, whatever happens first. We can simplify the analysis and remove the dependency from the random variable T , by assuming that W_s continues always for l steps and the state -1 is absorbing. Then, the lemma follows by just computing $E[W_l + 1]$. In particular, we have that

$$E[W_l + 1] = E[W_l + 1 | W_l \neq -1] \cdot \Pr[W_l \neq -1]. \quad (27)$$

Let \mathcal{T} be the step that W_t hits -1 for first time. Using the *Reflection Principle* we show can that for any nonnegative integer i it holds that

$$\Pr[\mathcal{T} = 2i + 1] = 2^{-(2i+1)} \frac{\binom{2i}{i}}{i + 1}. \quad (28)$$

It is direct that the W_t cannot be -1 for t even, i.e. $\Pr[\mathcal{T} = 2i] = 0$, for every positive integer i . It is direct to see that it holds that

$$\Pr[W_l = -1] = \Pr[\mathcal{T} \leq l] = 1 - \sum_{i > \lfloor (l-1)/2 \rfloor} 2^{-(2i+1)} \frac{\binom{2i}{i}}{i + 1}.$$

To this end we use Stirling approximation, i.e. for a sufficiently large n it holds that $n! = \sqrt{2\pi n} \left(\frac{n}{e}\right)^n e^{\lambda_n}$, with $\frac{1}{12n+1} \leq \lambda_n \leq \frac{1}{12n}$. Then we have that

$$\sum_{i > \lfloor (l-1)/2 \rfloor} 2^{-(2i+1)} \frac{\binom{2i}{i}}{i + 1} \leq \frac{1}{2\sqrt{\pi}} \sum_{i > \lfloor (l-1)/2 \rfloor} \frac{1}{i^{3/2}} \leq \sqrt{\frac{2}{\pi l}}.$$

Thus, we get that

$$\Pr[W_l = -1] \geq 1 - \sqrt{\frac{2}{\pi l}}. \quad (29)$$

On the other hand, it is direct to see that given that the walk W_t does not hit -1 it is just a random walk on the positive integers and it is a folklore result that

$$E[Z_l | Z_l \neq -1] \leq \sqrt{\left(\frac{2}{\pi}\right)} \cdot \left(1 + \frac{3}{2l}\right). \quad (30)$$

The lemma follows by plugging (29) and (30) into (27) and taking $l \geq 10$. \square

Proof of Proposition 4: Consider the revelation of the values of the matrix \mathcal{S} we gave above. Let t_i be the index of the column we have revealed up to round i . I.e. at round $i + 1$ we check whether $X_{t_i} = \sum_{j=1}^{t_i} \mathcal{S}(1, j) - \mathcal{S}(2, j)$ is zero or not. Let l the maximum number of columns in each submatrix R_j^i be equal to 10.

Given X_{t_i} and assuming that the coupling continuous, i.e. t_i the number of columns we have revealed so far is much smaller than N , we show that it holds that

$$E[X_{t_{i+1}} | X_{t_i}] \leq \frac{2.3}{\pi} X_{t_i}. \quad (31)$$

However, before showing the above let us see which are its consequences. Taking the average from both sides, we get

$$E[X_{t_i}] \leq \frac{2}{\pi} E[X_{t_{i-1}}] \leq \left(\frac{2.3}{\pi}\right)^i,$$

since $X_{t_1} = 1$ (it always holds that $\mathcal{S}(1, 1) - \mathcal{S}(2, 1) = 1$). It is also direct to see that it always holds that $X_{t_i} \leq l \cdot X_{t_{i-1}} \leq l^i$. That is, in round i we will need to reveal at most l^i columns of the matrix. This fact implies that the maximum number of rounds j we can have satisfies the condition that $\sum_{t=0}^j l^t \leq N$. Direct calculations suggest that $j \leq \frac{99}{100} \frac{\ln N}{\ln l} = 0.43 \ln N$, since $l = 10$. Clearly, the proposition follows once we show (31). For this we are going to use Lemma 9. Notice that given that at round i we have $X_{t_i} = \left| \sum_{j=1}^{t_i} (\mathcal{S}(1, j) - \mathcal{S}(2, j)) \right|$, for X_{i+1} the following holds:

$$X_{t_{i+1}} = \sum_{s=0}^{X_{t_i}} \left(1 + \sum_{j=1}^{T_s} R_s^i(1, j) - R_s^i(2, j) \right),$$

where T_s is the length of the submatrix R_s^i . From Lemma 9 we have that for any i, s it holds

$$E \left[1 + \sum_{j=1}^{T_s} R_s^i(1, j) - R_s^i(2, j) \right] \leq \frac{2.3}{\pi}.$$

Combining the above two relations and by linearity of expectation we get that

$$E[X_{t_{i+1}} | X_{t_i}] = \sum_{s=1}^{X_{t_i}} E \left[1 + \sum_{j=1}^{T_s} R_s^i(1, j) - R_s^i(2, j) \right] \leq \frac{2.3}{\pi} X_{t_i}.$$

The proposition follows. \square

5 Rest of the proofs

5.1 Proof of Lemma 1

Consider some integer k and let $\nu_{c_1} : [k]^d \rightarrow [0, 1]$, for some $c_1 \in [k]$, be the distribution over the d -dimensional vector induced by the following process: A vector S is distributed as in ν_{c_1} if component $S(i)$ is distributed uniformly at random in $[k] \setminus \{c_1\}$, independently of the other components, for every $i \in [d]$.

Observe that the information we have for L_X^j and L_Y^i is the following: For L_X^j we know that the colour $c \in L_X^j$, $q \notin L_X^j$ since $L_X(j) = q$, and there is at least one extra colour in $[k] \setminus \{c, q\}$ that does not appear in L_X^j . As far as L_Y^i is regarded, we know that $c \in L_Y^i$, $q \notin L_Y^i$ and $L_Y(i)$ is equal to a colour that does not appear in L_X^j .

So as to show the lemma it suffices to show that conditional the colouring of $L_Y(i)$, the distribution of L_X^j is identical to the one of L_Y^i . Assume that $L_Y(i) = s$, for some $s \in [k] \setminus \{c, q\}$.

Let the event $A = "L_Y^i \text{ is } j\text{-good and } L_Y(i) = s"$. For any $S \in [k]^d$ it holds that

$$Pr[L_Y^i = S | A] = \nu_s(S | B),$$

where $B = "there exists $t \in [d]$ such that $S(t) = c$ and there is no $t \in [d]$ such that $S(t) = q"$. Let Q be the number of colours in $[k] \setminus \{c, q\}$ that do not appear in L_X^j . It suffices to show that it holds that$

$$Pr[L_X^j = S | s, q \notin L_X^j, c \in L_X^j, Q > 0] = \nu_s(S | B). \quad (32)$$

Clearly we have that

$$\begin{aligned}
Pr[L_X^j = S | s, q \notin L_X^j, c \in L_X^j, Q > 0] &= \frac{Pr[L_X^j = S, Q > 0, s, q \notin L_X^j, c \in L_X^j]}{Pr[Q > 0, s, q \notin L_X^j, c \in L_X^j]} \\
&= \frac{Pr[s, q \notin L_X^j, L_X^j = S, c \in L_X^j]}{Pr[s, q \notin L_X^j, c \in L_X^j]} \\
&= Pr[L_X^j = S | s, q \notin L_X^j, c \in L_X^j]. \tag{33}
\end{aligned}$$

In the penultimate derivation we eliminated the event $Q > 0$ from both probability terms, in the nominator and denominator, since whenever $s \notin L_X^j$ holds it also holds that $Q > 0$. Then, it is straightforward that the r.h.s. of (33) is equal to $\nu_s(S|B)$, i.e. (32) holds as promised.

5.2 Proof of Lemma 2

The lemma follows by just examining the information we have for each of the four lists. As far as the i -good pair (L_X^t, L_Y^t) is concerned we have the following: $L_X(t)$ is distributed uniformly at random among the colours $[k] \setminus \{c, q\}$ that do not appear in L_X^i while $c \notin L_X^t$ and $q \in L_X^t$. Also, $L_Y(t)$ is distributed uniformly at random among the colours $[k] \setminus \{c, q\}$ that do not appear in L_Y^i while $q \notin L_Y^t$ and $c \in L_Y^t$.

As far as the i -fail pair (L_X^s, L_Y^s) is concerned we have the following: $L_X(s)$ is distributed uniformly at random among the colours $[k] \setminus \{c, q\}$ that do not appear in L_X^i while $q \notin L_X^s$ and $c \in L_X^s$. Additionally, $L_Y(s)$ is distributed uniformly at random among the colours $[k] \setminus \{c, q\}$ that do not appear in L_Y^i while $c \notin L_Y^s$ and $q \in L_Y^s$.

The lemma follows by noting that we have coupled the rescuable pair L_X^i and L_Y^i such that the colours in $[k] \setminus \{c, q\}$ that do not appear in both lists are exactly the same. Thus, we can couple identically $L_X(t)$ with $L_Y(s)$ and $L_X(s)$ with $L_Y(t)$. Then, it is direct that we can couple identically L_X^t with L_Y^s and L_X^s with L_Y^t .

6 Full Coupling

Coupling: $(X(v), Y(v))$

IF $X(v) = Y(v)$, then couple indentially the children of v .

For each w , child of v execute **Coupling** $(X(w), Y(w))$.

ELSE do the following:

Phase 1:- Partial revelation of the lists.

1. Reveal only for which i we have $L_X(i), L_Y(i) \in \{c, q\}$. Couple the choices of L_X and L_Y such that if $L_X(i) = q$, then $L_Y(i) = c$.
2. For each i such that $L_X(i) = q$ and $L_Y(i) = c$ reveal whether L_X^i and L_Y^i is “bad” or not. We use coupling such that the result of revelation to be the same for both L_X^i and L_Y^i .
3. For each pair of bad lists (L_X^i, L_Y^i) reveal whether they are “rescuable”. We use coupling such that the result of revelation to be the same for both L_X^i and L_Y^i . The coupling is so that *the set of colours in $[k] \setminus \{c, q\}$ that are not used to be the same for both rescuable lists L_X^i, L_Y^i .*
4. Partition the non-bad L_X^j s and L_Y^j s to (maximal) equally sized sets. The partitions are as many as the rescuable pairs and each rescuable pair is associated to exactly one partition. That is, the rescuable pair (L_X^i, L_Y^i) is associated to a set of indices $A_i \subseteq [d]$ such that the following holds: For any $t \in A_i$ L_X^t and L_Y^t belong to partition associated to (L_X^i, L_Y^i) .

5. For each i and for each $j \in A_i$ we reveal if the pair (L_X^j, L_Y^j) consists of i -special lists. We use coupling such that either both lists in the pair are i -special or both are not.

Phase 2: - List Association.

For each rescuable pair (L_X^i, L_Y^i) do the following:

1. Reveal each i -special pairs in A_i whether it is i -good or i -fail until either of the following two happens:
 - the number of i -good pairs exceeds the number of i -fails by one,
 - there are no other i -special pairs in A_i to reveal.
2. Reveal the remaining unrevealed i -special pairs, if any, by using identity coupling.
3. Match every i -good pair with one i -fail such that no two i -good pairs are matched to the same i -fail pair.
4. If the i -good pair (L_X^j, L_Y^j) is matched with the i -fail (L_X^s, L_Y^s) , then set $f(j) = s$ and $f(s) = j$.
5. If there is an i -good pair (L_X^j, L_Y^j) for which there is no i -fail to be matched, match it with the rescuable pair (L_X^i, L_Y^i) . Also, set $f(i) = j$ and $f(j) = i$.
6. For each $j \in A_i$ such that (L_X^j, L_Y^j) is not matched yet, match it to itself and set $f(j) = j$.

Phase 3: - Full revelation.

1. For every s, t such that $f(s) = t$, couple optimally $L_X(s)$ with $L_Y(t)$ as well as L_X^s with L_Y^t .
 2. Reveal which element of the list L_X is assigned to which child of r and which element of L_X^j goes to which grandchild of r , as Remark 6 specifies.
 3. Assuming that v , child of r , is such that $X(v)$ is set $L_X(s)$, then we set $Y(v)$ equal to $L_Y(t)$, where $t = f(s)$. Also, for u , child of v , such that $X(u)$ set $L_X^s(j)$ we set $Y(u)$ equal to $L_Y^t(j)$.
-

References

- [1] D. Achlioptas, A. Coja-Oghlan. *Algorithmic Barriers from Phase Transitions*. In Proc. of 49th Annual IEEE Symposium on Foundations of Computer Science, FOCS 2008.
- [2] D. Achlioptas and A. Naor. *The two possible values of the chromatic number of a random graph*. Annals of Mathematics, 162 (3), (2005), 1333-1349.
- [3] D. Aldous. *Random walks of finite groups and rapidly mixing Markov chains*. In: Séminaire de Probabilités XVII 1981/82, Springer-Verlag, Berlin. pp. 243-297.
- [4] N. Bhatnagar, J. Vera, E. Vigoda and D. Weitz. *Reconstruction for coloring on Trees*. SIAM J. Discrete Math. Vol 25, No. 2, pp 809-826, (2011).
- [5] C. Daskalakis, E. Mossel and S. Roch. *Optimal Phylogenetic Reconstruction*. In Proc. of the 38th Annual ACM Symposium on Theory of Computing (STOC), 159-168, 2006.
- [6] M. Dyer, A. Flaxman, A. M. Frieze and E. Vigoda. *Random colouring sparse random graphs with fewer colours than the maximum degree*. Random Struct. and Algorithms **29** (2006), pp. 450-465.

- [7] C. Efthymiou and P. G. Spirakis. *Random sampling of colourings of sparse random graphs with a constant number of colours*. In Theoretical Computer Science 407 (2008), pp. 134-154.
- [8] W. Evan, C. Kenyon, Y. Peres and L. J. Schulman. *Broadcasting on trees, the Ising model*. The Annals of Applied Probability, 2000, Vol. 10 No. 2, , 410-433
- [9] C. Efthymiou. *A simple algorithm for random colouring $G(n, d/n)$ using $(2+\epsilon)d$ colours*. In Proc. of the 23rd Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2012.
- [10] A. Gerschenfeld and A. Montanari. *Reconstruction for models on random graphs*. In Proc. of 48th Annual IEEE Symposium on Foundations of Computer Science, FOCS 2007.
- [11] H. O. Georgii. *Gibbs Measures and Phase Transitions*, de Gruyter Stud. Math. 9, de Gruyter, Berlin, 1988
- [12] H. Kesten and B. H. Stigum. *Additional limit theorems for indecomposable multidimensional Galton-Watson processes*. Ann. Math. Statist, 1966, 37 pp 1463-1481.
- [13] F. Krzakala, A. Montanari, F. Ricci-Tersenghi, G. Semerjian, L. Zdeborova. *Gibbs states and the set of solutions of random constraint satisfaction problems*. In Proc. National Academy of Sciences 104 (2007) pp 10318-10323.
- [14] A. Montanari, R. Restrepo and P. Tetali. *Reconstruction and Clustering in Random Constraint Satisfaction Problems*. SIAM J. Discrete Math. 25 (2): 771-808 (2011).
- [15] E. Mossel. *Reconstruction on Trees: Beating the second eigenvalue*. The Annals of Applied Probability 2001, Vol. 11, No. 1 285-300.
- [16] E. Mossel. *Phase transitions in phylogeny*. Trans. Amer. Math. Soc., 356(6):2379-2404, 2004
- [17] E. Mossel and Y. Peres. *Information flow on trees*. The Annals of Applied Probability 2003, Vol. 13, No. 3, 817-844.
- [18] E. Mossel and A. Sly. *Gibbs Rapidly Samples Colorings of $G_{n,d/n}$* . In journal Probability Theory and Related Fields, Vol. **148** (2010).
- [19] G. Semerjian. *On the freezing of variables in random constraint satisfaction problems*. J. Stat. Phys, 2008, Vol. 130, No 2, 251-293.
- [20] A. Sly. *Reconstruction of Random Colourings*, Commun. Math. Phys., 188 (2009), pp 943-961.